

**Shotcut video editing app and preattentive processing**

Suzanne Hillman

HF700: Foundations in Human Factors

Dr. Bill Gribbins

10/29/2023

## **Shotcut video editing app and preattentive processing**

Human perceptual systems - and therefore humans - are pattern detectors. Our visual processing system collects together and feeds forward relevant information to create feature maps in the visual cortex which allow us to detect specific types of visual patterns within approximately 250ms, or the amount of time it typically takes to start a new visual saccade (Land, 1999; Treisman & Gelade, 1980). These feature maps allow us to function in a world with far too much visual information to take in without processing and refinement. If we could not react quickly to potential threats, correctly identify objects that could be food, or navigate our surroundings, our ancestors would have perished long before they could reproduce.

Feature maps allow us to separate objects from each other in large part due to the fact that we combine features and pass them up as combinations, rather than as individual unconnected pieces of information (Livingstone & Hubel, 1988; Rosenholtz, 2020; Treisman & Gelade, 1980). This greatly improves the speed with which we can process and react to whatever those objects turn out to be, and allows for such behaviors as shying away from things that look like snakes before we consciously realize what we are looking at (Ohman et al., 2001).

Within the context of digital interfaces, this means that we detect patterns within an interface and rely on them to structure our understanding of it. If these patterns are missing, or applied incorrectly or poorly, we become anxious and confused without the unconscious cues that provide us with the context we need. The digital interface example I shall be exploring in this review is a video editor called [Shotcut](#).

### **Neurological underpinnings of pattern detection**

#### **Within the retina**

Pattern detection is the result of many vision cells working together to process, enhance, and feed forward information up the next step in the chain. This results in information that becomes increasingly more specialized and is fed forward more efficiently and faster than if these cells were not refining visual information (Wolfe & Horowitz, 2017).

Horizontal cells inhibit responses around bright spots in our visual field to make edges more noticeable and feed that information forward to the bipolar cells, which in turn is sent to the retinal ganglion cells (RGCs) to help create a center-surround receptive field to compare related features against each other (Chaya et al., 2017; Thoreson & Mangel, 2012).

Bipolar cells begin the process of forming patterns based on hue and the recent start or stop (transient) or ongoing (sustained) nature of light stimulus. These cells are separated into those that excite in response to a feature they are looking for - ON cells - and the ones that inhibit based on a feature they are looking for - OFF cells (Demb & Singer, 2015; Gollisch & Meister, 2010; Wässle, 2004). They feed forward processed and refined information to amacrine cells and RGCs to be integrated, enhanced, and fed forward to the lateral geniculate nucleus (LGN) which creates a retinotopic map and processes,

enhances, and feeds information forward to the visual cortex and develop feature maps (Demb & Singer, 2015; Wässle, 2004).

While amacrine cells are not well understood, we know that some mediate between the rod and cone pathways to incorporate rod information with cone information, others support motion sensitivity, still others shape the behavior of the inhibitory surrounds to support RGC center-surround receptive field formation, and others pass along direction-sensitive light responses (Demb, 2007; Demb & Singer 2012; Demb & Singer 2015; Sanes & Masland, 2015). Amacrine cells generally modulate and refine information from the bipolar cells before sending it along to the RGCs.

RGCs are the first cells in the vision pathway to have center-surround receptive fields for comparison of inputs based on various features. They integrate, maintain, and feed forward transient and sustained light information, changes in relative amounts of light within their receptive fields, luminosity levels, spatiotemporal changes in luminosity, hues, and direction sensitivity (Callaway, 2005; Demb, 2007; Dhande et al., 2015; Gollisch & Meister, 2010; Sanes & Masland, 2015).

### **Lateral Geniculate Nucleus**

Data from the RGCs are fed forward through the optical nerve to the LGN. The 6 layers of the LGN are retinotopically aligned, meaning a line drawn perpendicularly through the layers intersects with the same area of our vision in every layer, and this alignment largely matches our eyes' field of view, albeit with increased surface area for the foveal area. The retinoscopic map is fed forward to the visual cortex and may be important for the serial integration of features as well as localizing various feature maps (Dhande et al., 2015; Livingstone & Hubel, 1988).

Four of the layers of the LGN are the parvocellular (parvo) subdivision, which integrates, maintains, and feeds forward hue information to the visual cortex to support feature maps for hue (Callaway, 2005; Ghodrati et al., 2017; Livingstone & Hubel, 1988; Sanes & Masland, 2015).

The other two layers are the magnocellular (magno) subdivision, which refines, maintains, and feeds forward spatial information and movement selectiveness to the visual cortex to support creating feature maps for depth, spatial organization, and movement (Callaway, 2005; Ghodrati et al., 2017; Livingstone & Hubel, 1988).

Both subdivisions refine and feed forward information about orientation and shape to the visual cortex (Ghodrati et al., 2017; Livingstone & Hubel, 1988).

### **Visual Cortex**

Cells in V1 have subdivisions that select for hue, depth, movement, direction of movement, and specific orientation which are used to create feature maps for each of these dimensions (Livingstone & Hubel, 1988; Treisman & Gelade, 1980; Wandell et al., 2005).

#### ***Parvo network***

Parvo network data are sent to multiple layers of V1: 4C $\beta$ , Blob cells, and interblob cells. Cells in 4C $\beta$  process hue and orientation and contribute to the related feature maps (Livingstone & Hubel, 1988; Treisman & Gelade, 1980; Wandell et al., 2005; Yen & Finkel, 1998).

Blob cells process hue and luminance, with hue being selected for using double-opponent center-surround receptive fields. Double-opponent center-surround receptive fields mean that instead of a receptive field exciting in the presence of red in the center and inhibiting in the presence of green in the surround for the red-green pathway (or vice versa), the center will excite to either red or green in the center, and hues that are not red or green will be inhibited in the surround. Blob cells contribute to feature maps for hues, lines, and edges (Livingstone & Hubel, 1988; Treisman & Gelade, 1980; Wandell et al., 2005; Yen & Finkel, 1998).

Interblob cells process and contribute to feature maps for orientation without direction, and likely use hue information to identify color contrast edges to support feature maps for orientation of lines and edges (Livingstone & Hubel, 1988; Treisman & Gelade, 1980; Wandell et al., 2005; Yen & Finkel, 1998).

Parvo network data goes to V2 directly and through the blob and interblob cells. The blob cell data are fed forward to thin stripes in V2, which contribute to feature maps for hue. Interblob cell data go to pale stripes in V2, which contribute to feature maps for orientation. About half of the cells in the pale stripes in V2 are end-stopped, meaning that they fire the most for smaller stimuli and decrease their firing strength with larger stimuli. They are also likely reactive to hue contrast like the interblob cells (Livingstone & Hubel, 1988; Treisman & Gelade, 1980; Wandell et al., 2005; Yen & Finkel, 1998).

Finally, both pale and thin stripes feed data forward to V4, which supports feature maps for hue and orientation (Livingstone & Hubel, 1988; Treisman & Gelade, 1980; Wandell et al., 2005; Yen & Finkel, 1998).

### ***Magno network***

Magno network data is fed forward to 4C $\alpha$  in V1, which supports orientation feature maps. From there, the information is fed forward to 4B in V1, which supports feature maps for orientation and direction of movement. Next, it is fed forward to thick stripes in V2, which support feature maps for orientation and depth (Livingstone & Hubel, 1988; Treisman & Gelade, 1980; Wandell et al., 2005).

### **Pattern matching as it relates to digital interfaces**

We have now moved from points of light hitting various photoreceptors to feature maps for orientation, movement, direction of movement, hue, edges and boundaries, and depth (Wandell et al., 2005; Wolfe & Horowitz, 2017; Yen & Finkel, 1998). I shall now explain how these relate to the preattentive patterns that we detect and how this relates to digital interfaces.

Generally, our brains seem to use feature maps to group things that might be objects at a preattentive level, resulting in the automatic pattern matching we see (Wolfe & Utochkin, 2004; Wolfe &

Horowitz, 2019). Of course, as with the rest of the neurological system involved, there is variability within those feature maps based on the fact that the photoreceptors themselves and every step along the way have a reaction to a set of features with variations in the strength of the reaction. As a result, preattentive vision does poorly at discriminating stimuli features that are too similar to other stimuli on the same dimensions (Wolfe & Utochkin, 2004; Wolfe & Horowitz, 2019).

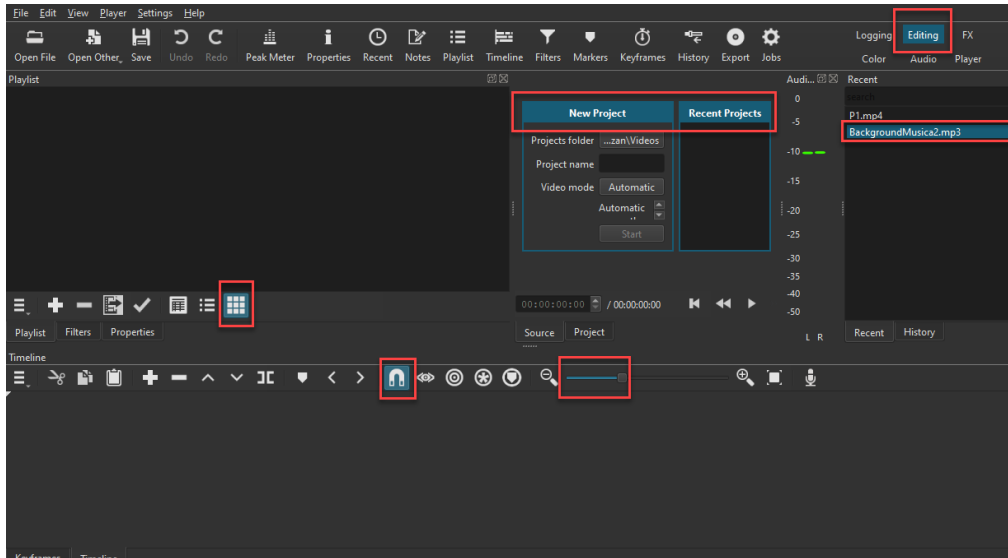
However, with sufficient difference between stimuli on the relevant dimensions that our feature maps include, preattentive processing is quite powerful. In most cases combining features in a stimuli changes them from parallel, preattentive processing to conscious, serial processing (Wolfe & Horowitz, 2004), but certain stimuli do not have this problem due to being entirely separate pathways. The best example of this is that of hue and spatial information - those are separated out into the parvo and magno streams in the LGN, and do not interfere with each other when combined (Wolfe & Utochkin, 2019).

### **Shotcut Video editing app**

The Shotcut video editing app in Figure 1 makes good use of closure around sections of the app to clarify that they are separate concepts. The concept of closure suggests a possible object based on edge feature maps that indicate an enclosed space. They have items that are the same hue of blue to suggest that they have a relationship to each other which I have enclosed in red in Figure 1. The same shade of blue suggests similarity to each other based on a hue feature map. Their buttons have a similar form to suggest a similarity relationship as well (Wolfe & Utochkin, 2019).

Most of the items of the same hue have the same meaning, but ‘new project’ and ‘recent projects’ are not selected items like the rest. Preattentive detection of similar items is evolutionarily important because it allows us to look at a berry bush ripe with berries, with each berry having approximately the same form, hue, and luminance, and correctly interpret them as similar to each other.

There’s clear alignment information in the horizontal lines separating the top buttons from the project information and the line above the timeline. However, there is no consistent alignment in the vertical direction, making it difficult to identify a pattern to follow and is therefore a high value thing to fix. Alignment patterns appear based on orientation feature maps suggesting the possibility of an object with the alignment line as a boundary of that object, and do a great deal toward helping us understand relationships between elements of an interface (Wolfe & Utochkin, 2019).

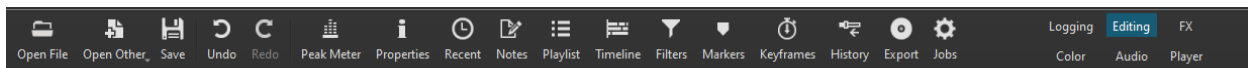


**Figure 1**

Shotcut interface without a project loaded in.

One of the most important things to fix, as seen in Figure 2, is that they do not separate any sets of buttons with whitespace or clearly visible edges. This makes the long line of buttons at the top look like they have a single purpose due to preattentive processing suggesting that things that are near each other could be a single object (Wolfe & Utochkin, 2019).

There is a very subtle line between the first three and the rest of the buttons, but it is only visible if one has greatly increased the size of the interface. Proximity relationships are implied for the other two sets of buttons, which also have excessively subtle lines between some of those buttons. The long list of buttons makes it very difficult to know where to start interacting with this tool. However, in Figure 2 there is good use of whitespace to suggest a relationship between the set of words off to the right, and a lack of relationship with the buttons on the left.

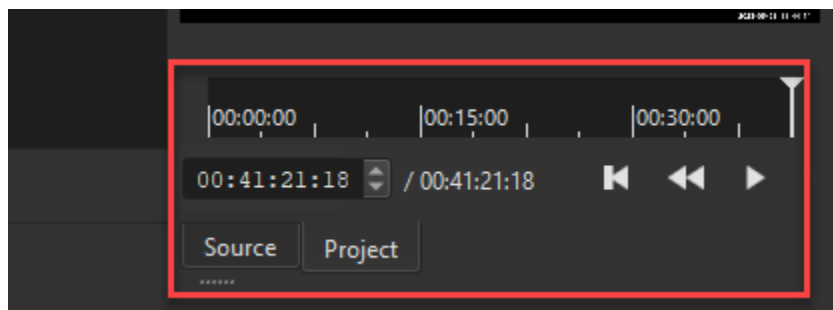


**Figure 2**

Closeup screenshot of the top-most navigation bar

In Figure 3, the timing scale, buttons for acting upon the video, and source and project appear related based on proximity and alignment. An incomplete vertical line from the current time in the project and the source, as well as between the end of the time scale and the play button serves to suggest alignment (Wolfe & Utochkin, 2019). Given that our eyes are accustomed to seeing objects with occlusions, it should come as no surprise that incomplete lines - often seen in alignment and grid

arrangements - are enough to be interpreted as a possible object. Unfortunately, the source and project labels are not related to the rest of the items.



**Figure 3**

Timing scale and source/project relationship

### Conclusion

Pattern detection has a strong basis in neurology from the very start of our perception through to a point where we detect certain patterns (mostly) preattentively. In the eyes, it is the visual cortex that combines the various features that previous cells have been organizing and passing along to come up with feature maps that serve as indicators of potential similarities in the world, whether a possible indication of an object or that of things that could be the same type of object. These patterns help us interpret the world fast enough to navigate within it, react to potential threats, and take advantage of potentially good things like berries.

Using these feature maps and their results in interfaces helps us make sense of the interface, just as these patterns help us make sense of the world, and their presence makes an interface easier and more comfortable to use.

## References

- Callaway, E. M. (2005). Structure and function of parallel pathways in the primate early visual system. *The Journal of Physiology*, *566*(1), 13–19.  
<https://doi.org/10.1113/jphysiol.2005.088047>
- Chaya, T., Matsumoto, A., Sugita, Y., Watanabe, S., Kuwahara, R., Tachibana, M., & Furukawa, T. (2017). Versatile functional roles of horizontal cells in the retinal circuit. *Scientific Reports*, *7*(1), 5540. <https://doi.org/10.1038/s41598-017-05543-2>
- Demb, J. B. (2007). Cellular Mechanisms for Direction Selectivity in the Retina. *Neuron*, *55*(2), 179–186. <https://doi.org/10.1016/j.neuron.2007.07.001>
- Demb, J. B., & Singer, J. H. (2012). Intrinsic properties and functional circuitry of the AII amacrine cell. *Visual Neuroscience*, *29*(1), 51–60. <https://doi.org/10.1017/S0952523811000368>
- Demb, J. B., & Singer, J. H. (2015). Functional Circuitry of the Retina. *Annual Review of Vision Science*, *1*(1), 263–289. <https://doi.org/10.1146/annurev-vision-082114-035334>
- Dhande, O. S., Stafford, B. K., Lim, J.-H. A., & Huberman, A. D. (2015). Contributions of Retinal Ganglion Cells to Subcortical Visual Processing and Behaviors. *Annual Review of Vision Science*, *1*(1), 291–328. <https://doi.org/10.1146/annurev-vision-082114-035502>
- Ghodrati, M., Khaligh-Razavi, S.-M., & Lehky, S. R. (2017). Towards building a more complex view of the lateral geniculate nucleus: Recent advances in understanding its role. *Progress in Neurobiology*, *156*, 214–255. <https://doi.org/10.1016/j.pneurobio.2017.06.002>
- Land, M. F. (1999). Motion and vision: Why animals move their eyes. *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology*, *185*(4), 341–352.  
<https://doi.org/10.1007/s003590050393>
- Livingstone, M., & Hubel, D. (1988). Segregation of Form, Color, Movement, and Depth: Anatomy, Physiology, and Perception. *Science*, *240*(4853), 740–749.  
<https://doi.org/10.1126/science.3283936>
- Ohman, A., Flykt, A., & Esteves, F. (2001). Emotion Drives Attention: Detecting the Snake in the Grass. *Journal of Experimental Psychology: General*, *130*(3), 466–478.  
<https://doi.org/10.1037/AXJ96-3445.130.3.466>



- Rosenholtz, R. (2020). What Modern Vision Science Reveals About the Awareness Puzzle: Summary-statistic encoding plus decision limits underlie the richness of visual perception and its quirky failures. *Atten Percept Psychophys*.
- Sanes, J. R., & Masland, R. H. (2015). The Types of Retinal Ganglion Cells: Current Status and Implications for Neuronal Classification. *Annual Review of Neuroscience*, 38(1), 221–246. <https://doi.org/10.1146/annurev-neuro-071714-034120>
- Thoreson, W. B., & Mangel, S. C. (2012). Lateral interactions in the outer retina. *Progress in Retinal and Eye Research*, 31(5), 407–441. <https://doi.org/10.1016/j.preteyeres.2012.04.003>
- Treisman, A. M., & Gelade, G. (1980). A Feature-Integration Theory of Attention. *Cognitive Psychology*, 12, 97–136.
- Wässle, H. (2004). Parallel processing in the mammalian retina. *Nature Reviews Neuroscience*, 5(10), 747–757. <https://doi.org/10.1038/nrn1497>
- Wandell, B. A., Brewer, A. A., & Dougherty, R. F. (2005). Visual field map clusters in human cortex. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1456), 693–707. <https://doi.org/10.1098/rstb.2005.1628>
- Wolfe, J. M., & Utochkin, I. S. (2019). What is a preattentive feature? *Current Opinion in Psychology*, 29, 19–26. <https://doi.org/10.1016/j.copsyc.2018.11.005>
- Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, 5(6), 495–501. <https://doi.org/10.1038/nrn1411>
- Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, 1(3), 0058. <https://doi.org/10.1038/s41562-017-0058>
- Yen, S.-C., & Finkel, L. H. (1998). Extraction of perceptually salient contours by striate cortical networks. *Vision Research*, 38(5), 719–741. [https://doi.org/10.1016/S0042-6989\(97\)00197-1](https://doi.org/10.1016/S0042-6989(97)00197-1)